


RESEARCH

Open Access



Cross-platform comparison of precision and time effectiveness of automated versus semi-automated brain volumetric measurements in healthy Egyptian adults

Abdallah Adel Abdallah Hassan^{1*} , Amr Osama Azab¹, Habeba Mostafa Ahmed¹,
Lena Mohamed Abdel Monem Mohamed¹, Mohamed Adel² and Ayman A. El-Basmy¹

Abstract

Background The primary goal of this study is to evaluate the precision, time effectiveness and reproducibility of open source automated magnetic resonance imaging (MRI) brain volumetric measurement platforms as compared to the semi-automated methods. The secondary goal of the study is to create age-stratified database of MRI brain volumetric measures of healthy Egyptian adults.

Results Statistically significant variation were found between the absolute but not the relative volumetric measures (P -value < 0.001). Both the absolute and relative global volumetric measures were strongly correlated with age. Only the absolute volumes showed strong correlation with gender with male volumes being significantly higher than females (P -value < 0.0001), while the normative ratio correlation with gender varied with the utilized software.

Conclusions FMRIB's software library automated segmentation tool (FSL-FAST) was considered the most efficient program to run global volumetric analysis, unless concurrent advanced volumetric analysis is needed or gray matter volume is the main objective of the volumetric analysis, then FreeSurfer is considered the most efficient volumetric software.

Keywords Automated MRI brain volumetry, FSL-FAST, FSL-FIRST, FreeSurfer, CAT12-SPM, Global brain volumetry

Background

The expanding burden of neurodegenerative and neuro-inflammatory conditions mandate developing precise biomarkers for early disease diagnosis, follow-up of therapeutic response and prognosis. Magnetic resonance imaging (MRI) brain volumetric analysis is an evolving

accurate and reproducible biomarker that serves such a purpose. For example, tracing early hippocampal and entorhinal cortex atrophy serves as an early diagnostic biomarker of Alzheimer disease, likewise detection of global and subcortical volumetric brain atrophy; in particular, early reduction of thalamic volume has been considered as a marker to predict the evolution from clinically isolated syndrome to clinically definite Multiple Sclerosis (MS) [1, 2].

Manual and semi-automated brain volumetric measurements are considered the gold standard methods for accurate brain volumetric analysis. However, both are extremely time-consuming and require deep expertise with the anatomical details; therefore, their applications

*Correspondence:

Abdallah Adel Abdallah Hassan
abdullah.hassan37318@postgrad.kasralainy.edu.eg

¹ Diagnostic and Interventional Radiology, Kasr Alainy Faculty of Medicine, Cairo University, Cairo, Egypt

² Computer Engineering Department, Faculty of Engineering, Cairo University, Cairo, Egypt



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

in day to day clinical practice are limited. On the other hand, automated brain volumetric analysis has multiple potential advantages over manual and semi-automated segmentation such as significant processing time reduction, excellent reproducibility and waving the need for extensive training or anatomical knowledge for the investigator [3].

However, automated brain volumetric measurements possess potential disadvantages worthy of investigation such as the limited options to manually fix any improperly segmented structures as well as method-specific bias and algorithmic variances that could cause disagreement between the results of different automated brain volumetric packages and manual/semi-automated brain volumetric measures [4].

Multiple open source software packages for automated brain volumetric measurement including Statistical Parametric Mapping (SPM), Computational Anatomy Toolbox (CAT), FMRIB’s software library (FSL) and BrainSuite (University of California, Los Angeles (UCLA)) were heavily investigated in comparison with each other as well as manual segmentation, yet most of the research work focused on precision of results and disregarded the time efficiency and sufficiency for different clinical practice settings as well as the technical differences between software that might underlay the variations in their volumetric results [5].

The primary aim of the study was to evaluate the precision, time effectiveness and reproducibility of open source automated MRI brain volumetric measurement platforms as compared to the semi-automated methods. The secondary aim of the study was to create age-stratified database of MRI brain volumetric measures of healthy Egyptian adults.

Methods

Patients

This prospective study carried out from 2019 to 2021 for healthy Egyptian participants. Structural MR images (n=395) were obtained from healthy Egyptian adults from both genders and different age groups (ranging from 15: ≤ 64 years of age) with representative sample for each age group calculated according to 2019 count of Egyptian population (Table 1).

Inclusion criteria

Cases with following criteria were included: healthy Egyptians 15: ≤ 64 years old, from both genders with no systemic or neurologic complaints except for mild infrequent headache with no other related red flag signs.

Exclusion criteria

Cases with following criteria were excluded:

- 1 All specific vulnerable groups including children, pregnant women, mentally disabled persons, economically and educationally disadvantaged persons.
- 2 Given history of significant head trauma.
- 3 Given history of head and neck intervention.
- 4 Given history of general anesthesia.
- 5 Given history of any extracranial malignancy, immunotherapy, chemotherapy or radiotherapy.
- 6 Any systemic diseases, e.g., diabetes mellitus (DM), hypertension, renal or liver impairment.
- 7 Given history or current neurological complaints.

Table 1 Representative sample for each age group calculated according to 2019 count of Egyptian population

| Age groups | % to total population | % to targeted age group (adults from 15–64) | | | | Males | Females |
|------------|-----------------------|---|---|---------|---|-----------|-----------|
| | | Males | | Females | | | |
| | | % | Number in the sample volume (total 385) | % | Number in the sample volume (total 385) | | |
| 15–19 | 9.7 | 7.7 | 30 | 7.3 | 29 | 4,367,988 | 4,131,991 |
| 20–24 | 10.3 | 8.1 | 32 | 7.8 | 31 | 4,623,621 | 4,424,043 |
| 25–29 | 9.7 | 7.6 | 30 | 7.3 | 29 | 4,334,645 | 4,175,258 |
| 30–34 | 7.8 | 6.1 | 24 | 5.9 | 23 | 3,456,601 | 3,364,004 |
| 35–39 | 6.1 | 4.7 | 19 | 4.6 | 18 | 2,711,932 | 2,639,282 |
| 40–44 | 5.5 | 4.2 | 17 | 4.2 | 17 | 2,422,954 | 2,379,682 |
| 45–49 | 5.0 | 3.9 | 16 | 3.8 | 15 | 2,222,893 | 2,184,980 |
| 50–54 | 4.4 | 3.4 | 14 | 3.3 | 13 | 1,922,803 | 1,903,745 |
| 55–59 | 3.6 | 2.7 | 11 | 2.7 | 11 | 1,567,139 | 1,557,610 |
| 60–64 | 2.6 | 2 | 8 | 2 | 8 | 1,144,789 | 1,146,574 |

Methods

Informed consent was obtained from the all participants, and they were subjected to:

History taking

Age, gender, psycho-neurological or cognitive symptoms.

MRI data acquisition

MRI data acquisition: 3D T1 MPRAGE acquired on Siemens Aera 1.5 Tesla machines (Germany), 20 channel head coil. The examination parameters were as follows: Voxel size: $1.0 \times 1.0 \times 1.0 \text{ mm}^3$, Field-of-view: 250 mm^2 , Repetition time: 2200 ms, Inversion time: 900 ms, Echo time: 2.88 ms, and Flip angle: 8° .

2D Fluid-attenuated inversion recovery (FLAIR) images and diffusion-weighted images (DWI) were acquired to rule out any asymptomatic structural brain abnormalities.

MRI data preprocessing

Syngovia V10 workstation used to withdraw data from MRI scanner. Image analysis was done using in house workstation Dell Precision T5600. The acquired data set quality was double-checked by two neuroradiologists. Any image data set that failed to pass quality check was excluded from the study. The reviewed image data sets were anonymized and converted from DICOM to NIFTI (Neuroimaging Informatics Technology Initiative) format using "dcm2nii" software. Automated analysis of the global brain volumetric measurements (total intracranial volume, gray matter volume, white matter volume and cerebrospinal fluid (C.S.F) volume) was done using: Free surfer, FAST, FIRST, SPM-CAT (Computational Anatomy Toolbox).

Automated volumetric analysis of the MRI data

In-house processing code development: Due to the massive data set acquired, an in-house processing code was developed to do sequential auto-run of the segmentation steps in the following software across all subjects. Continuous monitoring of the analysis process and checking the error log files of the segmentation software for any error was done. Segmentation algorithms of the individual software:

1. *Free surfer volumetric analysis* All acquired MRI data were segmented using free surfer "recon-all" segmentation pipeline. The segmentation results were exported to "aseg.stats" file.
2. *FSL-FAST volumetric analysis* The brain extraction tool (BET) of the FSL was used to perform skull stripping (removing the non-brain tissue from the images). The extracted brain mask (skull stripped

data) was analyzed using "fast" segmentation pipeline. The volumetric quantification of the generated masks was done using "fslstats" function of the FSL package.

3. *FSL-FIRST volumetric analysis* All acquired MRI data were segmented using FIRST segmentation pipeline using "run_first_all" function for segmentation of the subcortical brain structures. The volumetric quantification of the generated masks was done using "fslstats" function which is part of FSL package used for volume measurements.
4. *SPM-CAT (Computational Anatomy Toolbox) volumetric analysis* All acquired MRI data were segmented using MATLAB2015a based SPM12-CAT toolbox. The generated global brain volumetric results were automatically exported to pdf formatted report.

Semi-automated brain volumetry

The alignment of the generated masks to the boundaries of the anatomical structures was reviewed for the need of manual editing (semi-automated brain volumetry), but neither of these masks required any significant modification, so semi-automated segmentation was not required in this study.

Postprocessing of the MRI data

In-house numerical volumetric data auto-extraction code development: Due to the massive volumetric data generated, an in-house extraction code was developed to auto-extract the data of all subjects to a single master excel sheet to avoid human error during the data entry step.

Statistical analysis

Statistical analysis was done using Microsoft Excel 2016 statistical package (Microsoft Corporation. (2016). Microsoft Excel. Retrieved from <https://office.microsoft.com/excel>).

The mean, standard deviation, minimum, and maximum were calculated for the absolute and relative GMV, WMV, CSFV as well as the TIV generated by each software.

The Pearson's correlation test was used to determine the relationship between overall quality score and total number of topological defects requiring fixation by Free-Surfer. The single-factor analysis of variance (ANOVA) test was used to assess statistically significant differences between the absolute GMV, WMV, CSFV and TIV values generated by all segmentation software. The t-test was used to assess statistically significant differences between the absolute GMV, WMV, CSFV and TIV values generated by all segmentation software between males and

females. The F-test was used to assess any statistically significant differences between the absolute and relative GMV, WMV, CSFV values generated by all segmentation software between males and females. A statistically significant *P* value was less than 0.05.

Results

No significant correlation found between the participant’s age (therefore age-related parenchymal changes that could distort image signal and disrupt the segmentation process such as basal ganglia calcifications, cortical iron deposition and many others) and the overall quality score or total number of topological defects/holes detected and automatically corrected by FreeSurfer with $r = -0.028$ and 0.016 , respectively.

The processing time required by FSL-FAST was 8.45:13.58 min to perform global brain segmentation [gray matter volume (GMV), white matter volume (WMV), cerebrospinal fluid volume (CSFV) and TIV only), while CAT12 required 26: 43 min to complete the same task. FSL-FIRST required 9.27: 14.48 min to complete subcortical gray matter segmentation. FreeSurfer required around 10.37: 13.41 h to perform full brain segmentation including various global, subcortical and gyral volumetry among other topological brain measures like cortical thickness and surface area as well as extra-axial structures volumetry such as optic chiasm.

Using single factor Analysis of Variance (ANOVA), the absolute GMV, WMV, CSFV and TIV values generated by one segmentation software (CAT12-SPM, FreeSurfer and FSL-FAST) showed statistically significant variation from the values generated by the others. The absolute CSFV (Fig. 1A) showed the highest variations, followed by GMV (Fig. 1B) and TIV (Fig. 1C), while WMV showed the least degree of variation (Fig. 1D). The F-value for CSFV, GMV, TIV and WMV was constantly higher than the F-critical being 131.08, 42.425, 13.547 and 6.196 respectively with corresponding *P*-values of 1.14E-50, 0, 1.6E-06 and 0.002124.

On the contrary, the GMV, WMV, CSFV normative ratios (% of TIV) did not show any statistical significant variations between the different software platforms (CAT12-SPM, FreeSurfer and FSL-FAST software). The F-value for CSFV, GMV and WMV normative ratios were constantly lower than the F critical being 0.604697, 0.376954 and 0.513551 with corresponding *P*-values of 0.546462, 0.686057 and 0.598543 respectively.

Using F-test, variations of TIV values in Egyptians were primarily modulated by age irrespective of the utilized software with p -value=0 for each of CAT12-, FreeSurfer- and FSL-FAST-generated TIV data set. Such a variation was better appreciated in males more than in females and in FSL-FAST more than in CAT12- and

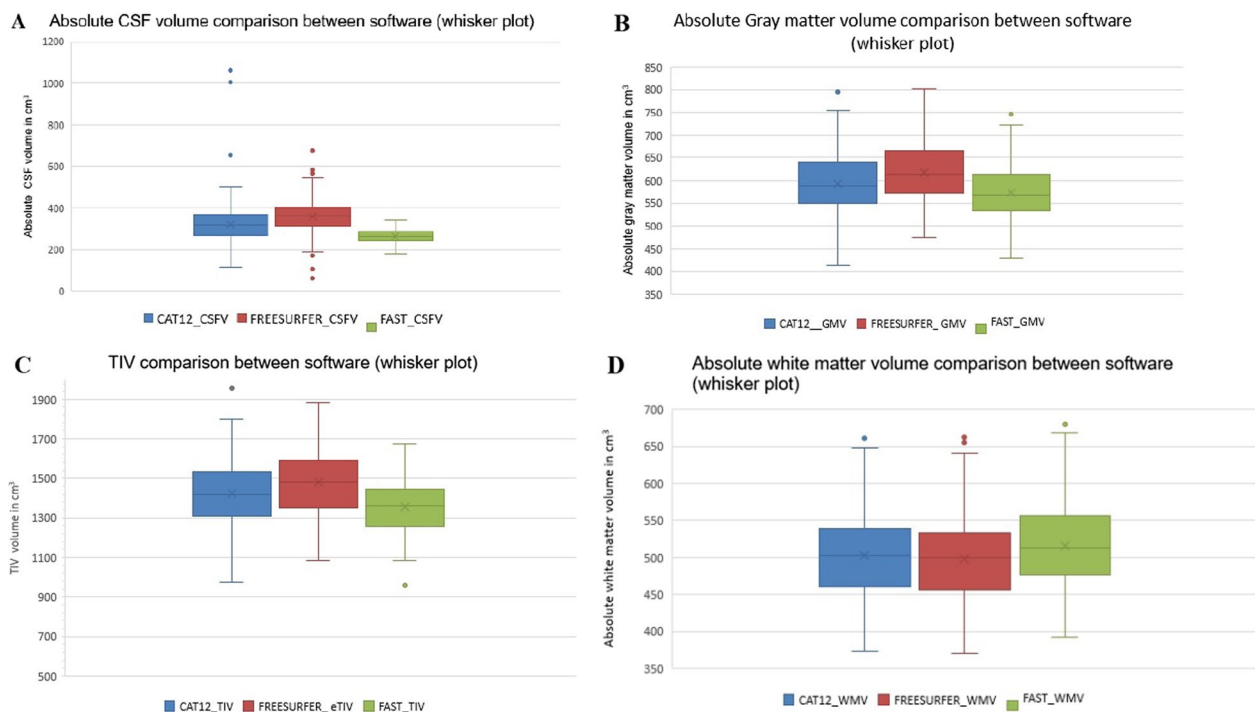


Fig. 1 A Whisker plot of the absolute CSFV values. B Whisker plot of the absolute GMV values. C Whisker plot of the absolute TIV values. D Whisker and line plots of the absolute WMV values generated by CAT12, FreeSurfer and FSL-FAST

FreeSurfer-generated data sets with FSL-FAST-generated male TIV showing $R^2=0.1$ (Fig. 2A, B).

Using F-test, both the absolute and the normative ratio of the GMV of adult Egyptians showed significant variation with age irrespective of gender and utilized software with p -value=0 for each of CAT12-, FreeSurfer- and FSL-FAST-generated absolute GMV values. P -values for GMV normative ratio generated by CAT12, FreeSurfer and FSL-FAST were $5.9464E-39$, $6.57E-64$ and $4.29421E-33$ in females and $5.56E-43$, $4.89247E-56$ and $2.45E-74$ in males respectively (Fig. 3A, B). Variation of GMV with age was more appreciated in males rather than females and FreeSurfer rather than CAT12- and FSL-FAST-generated data sets with FreeSurfer-generated male GMV showing $R^2=0.0768$.

FreeSurfer yielded the largest mean absolute GMV values and least degree of variance for both genders with average absolute GMV being 674.836 cm^3 for males and 578.135 cm^3 in females, followed by CAT12

which showed mean absolute values of 646.103 cm^3 and 554.083 cm^3 for males and females respectively yet with the highest degree of variance, FSL-FAST yielded the least mean absolute GMV values being 622.184 cm^3 and 536.174 cm^3 for males and females, respectively, yet with intermediate degree of variance.

Using two tailed t test; the average absolute GMV of Egyptian males was significantly higher than females irrespective of the utilized analysis software with p -value of $2.8E-46$, $7.82215E-59$ and $9.48E-47$ for CAT12-, FreeSurfer- and FSL-FAST-generated GMV, respectively (Fig. 4).

Only the FSL-FAST-generated WMV normative ratio conformed to the aforementioned statistical significant difference between absolute WMV in males and females with p -value of $2.5816E-05$, while both CAT12- and FreeSurfer-generated WMV normative ratios did not capture such a significant variation between male

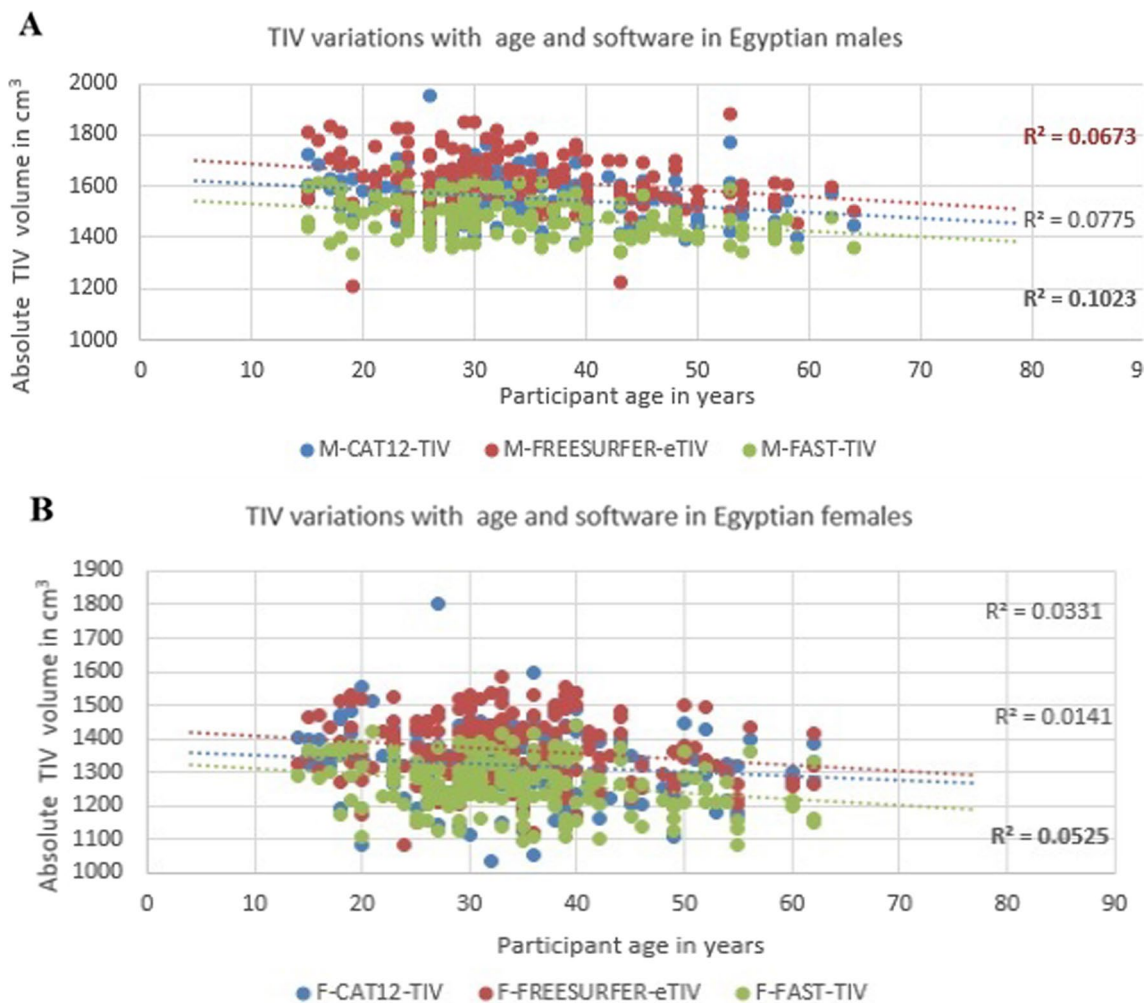


Fig. 2 Scatter plot of the TIV variations with Age and software in Egyptian males (A) and females (B)

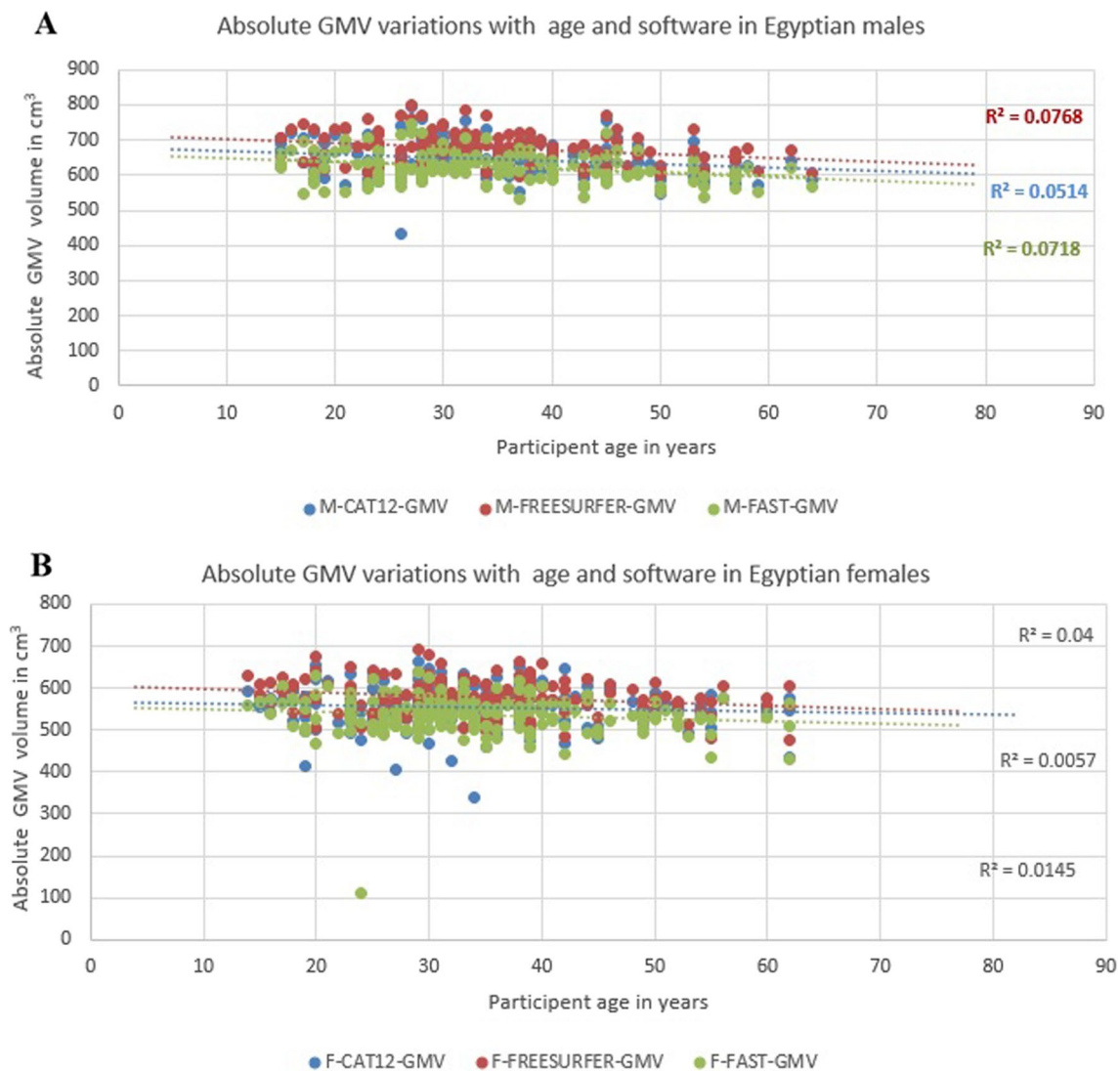


Fig. 3 Scatter plot of the Absolute GMV variations with Age, gender and software, M: male (A), F: female (B)

and female WMV with p -value of 0.272324088 and 0.997540, respectively.

Using F-test, both the absolute and the normative ratio of the CSFV of adult Egyptians showed significant variation with age irrespective of gender and utilized software with p -value=0 for each of CAT12-, FreeSurfer- and FSL-FAST-generated absolute CSFV values. P -values for CSFV normative ratio generated by CAT12, FreeSurfer and FSL-FAST were $1.63E-19$, $1.1E-45$ and $2.47E-68$ in females and $2.75E-29$, $3.57E-37$ and $1.77E-88$ in males, respectively (Fig. 5a, b). Variation of CSFV with age was more appreciated in males rather than females with FreeSurfer performing better in males showing an $R^2=0.0102$, and FSL-FAST performing better in females showing an $R^2=0.0114$.

Only the FSL-FAST-generated CSFV normative ratio conformed to the aforementioned statistical significant difference between absolute CSFV in males and females with p -value of 0.014867, while both CAT12- and FreeSurfer-generated CSFV normative ratios did not capture such a significant variation between male and female GMV with p -values of 0.066175 and 0.0640 58, respectively.

Discussion

In this study, we compared the efficiency and effectiveness of various automated brain segmentation software platforms, namely SPM-CAT12, FreeSurfer, FSL-FAST and FSL-FIRST in generating global brain volumetric results in terms of time consumption; user-friendly

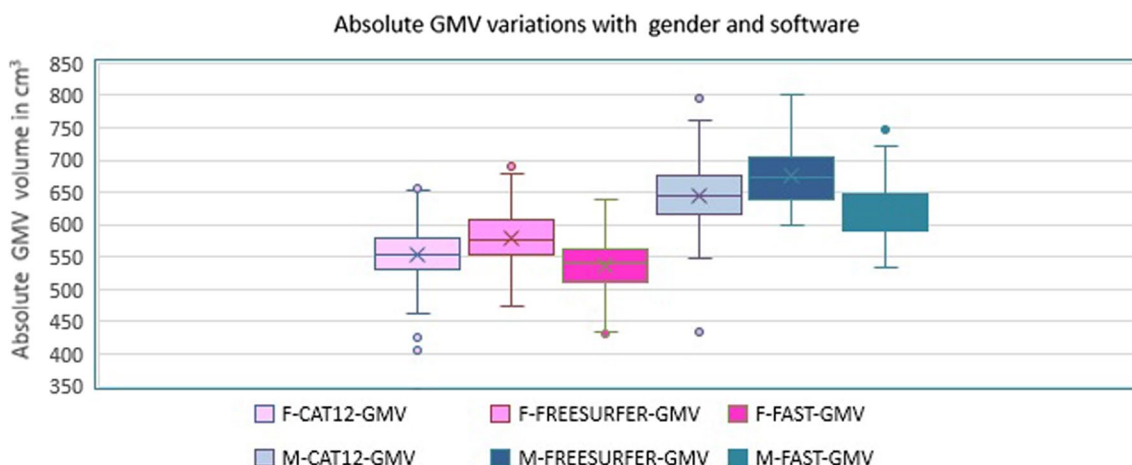


Fig. 4 Whisker and box plot of the Absolute GMV variations with gender and software. The rose, pink and magenta columns corresponds to volumetric analysis of female brain using CAT12, FreeSurfer and FAST respectively. The gray, blue and turquoise columns corresponds to volumetric analysis of male brain using CAT12, FreeSurfer and FAST respectively. The gray matter volume (GMV) of the male brain larger than the female (p-value of 2.8E-467, 82215E-59 and 9.48E-47 for CAT12-, FreeSurfer- and FSL-FAST-generated GMV, respectively). F: female, M: male

Graphical User Interface (GUI) and final numeric volumetric results; range of volumetric measures that can be possibly obtained by each software and its sufficiency to the different clinical scenarios; robustness of manual editing of the generated volume masks if required; inter-software consistency of the generated numeric volumetric results. The generated absolute and relative global volumetric values were used to calculate the first of a kind normative global volumetric data set for Egyptians.

524 Egyptians representing various age groups and both genders were enrolled, but only 463 participants passed the medical history/clinical assessment, 12 of them were discovered to be claustrophobic and could not complete the MRI scan. MRI data sets of 451 scans were extracted, anonymized, and converted to NIFTI format.

Two independent radiologists excluded 53 cases: 3 cases for motion and susceptibility artifacts designated R2 and R3 and 50 cases for incidental structural abnormality or pathological causes ($n=50$) with white matter foci being the most common cause of exclusion from the study accounting for 31 case (55.35% of all excluded cases).

Quantitative quality check using SPM-CAT12 excluded 3 other cases for failing to maintain the required image quality parameters showing an average quality score $\leq C+$. The final included MRI data set of 395 participants (201 females and 194 males) had an overall image quality score of B-, the selected scans had a minimal bias field (A-), extremely consistent image resolution (B+) and minor noise (B) compared to BWP data.

Due to the inherit limitations of each of the examined software platforms, several in-house python3-based

codes were developed to auto-run volumetric analysis across all subjects. The auto-generated volumetric masks were manually reviewed, yet neither of them required any significant manual editing.

Another set of in-house codes were developed to abolish any possible errors in extracting the targeted global volumetric numeric data from the massive statistical output files totaling 9085 file (23 output files per subject; 18 of them belong to the FreeSurfer with each of the FreeSurfer files containing over 316 numeric volumetric datum, and the rest of programs generate between 4:13 numeric volumetric data per file), the developed codes auto-extracted the targeted global volumetric numeric results and cross-tabulated them into a single Excel sheet, then total white matter and CSF volume of the FreeSurfer were manually calculated using Backhausen et al. [6] method, and statistical analysis was carried out using Microsoft Excel 2016 statistical package.

We found that FSL-FAST consumed the least processing time to calculate the absolute global brain volumetric measures, yet it does not have a user-friendly GUI, manual editing interface, and required multiple code development to export its results into.txt format and to run multi-subject analysis sequentially.

SPM-CAT12 consumed triple the time required by FSL-FAST, but it performed background image quality analysis; possessed user-friendly GUI and output file format; allowed selecting from variable range of basic to advanced volumetric analysis; auto-calculated the normative ratio along with absolute measures; allowed limited manual editing of its volumetric segments through same interface.

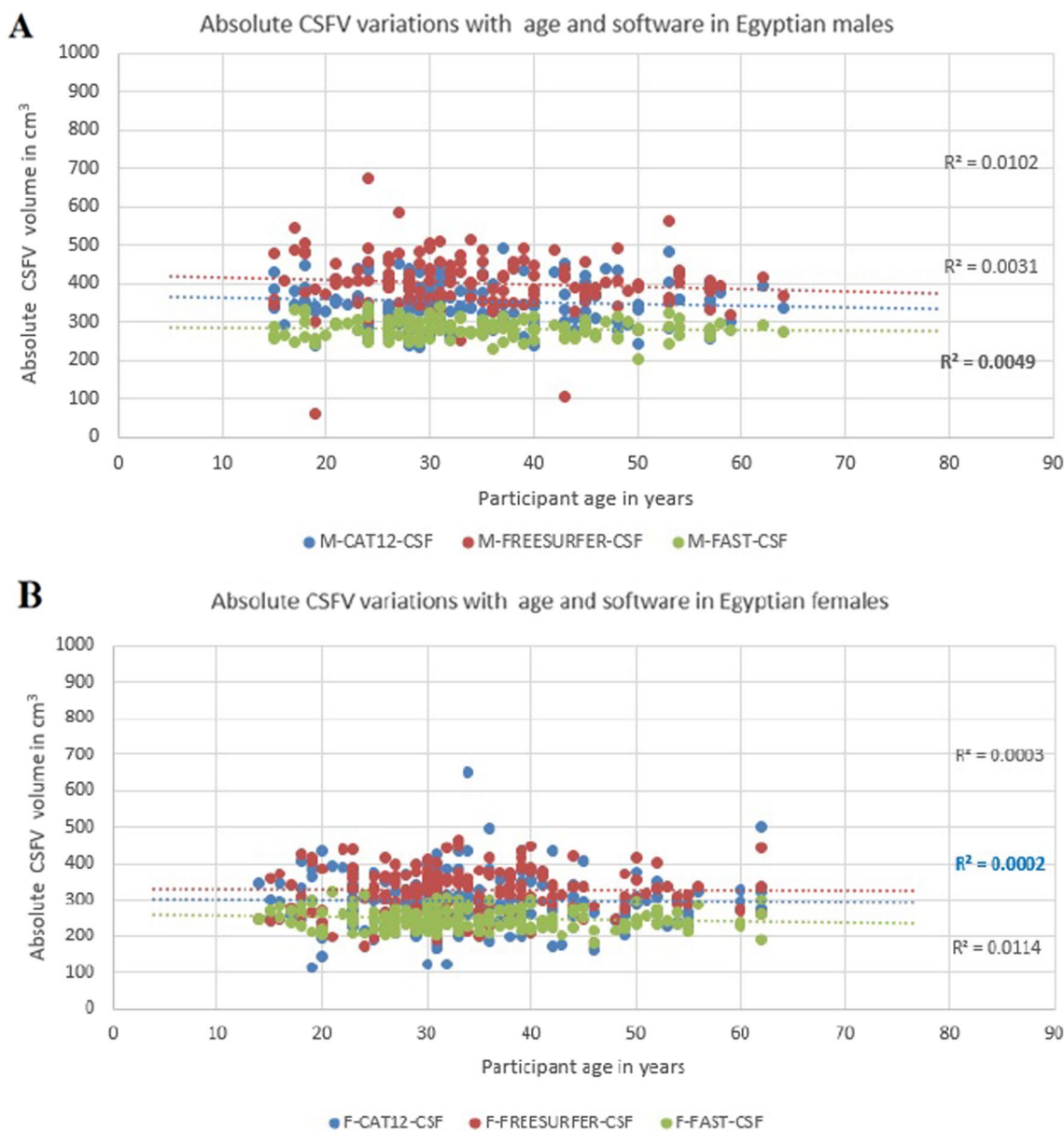


Fig. 5 Scatter plot of the absolute CSFV variations with age, gender and software, M: male (A), F: female (B)

FreeSurfer generated full range basic to advanced subsegmental volumetric analysis without allowing the user to choose certain range of analysis; therefore, it consumed the longest processing time being 25 times that of SPM-CAT12 and produced massive volume of structural numerical measures (20 files, most of them contains >300 numerical volumetric datum) along with white matter hypo-intensities volume. FreeSurfer requires high processing power, does not have a user-friendly GUI, yet, running the analysis is fairly easy and full range of manual editing can be performed through

its viewer. FreeSurfer’s Image quality assessment tools are still under development.

Neither of these programs required significant manual editing of their masks especially the FreeSurfer as in McCarthy et al. [7], and Beelen et al. [8], works therefore, semi-automated volumetric analysis was waived for this study.

In this study, we found that age-related brain changes like hemosiderosis did not affect the required analysis time; on the contrary, the overall average image quality of the included MRI scans inversely correlated

with the analysis time (the higher the image quality, the lesser the number of surface defects and therefore the lesser the processing time required by FreeSurfer). Such a finding was unparalleled in literature and may worth further investigation and quantification to assess its possible application in increasing time efficiency of FreeSurfer, especially in light of Fellhauer et al. [9], which linked the accuracy of volumetric analysis to image quality.

In this study, the absolute global volumetric measures (GMV, WMV, CSFV and TIV) varied significantly from one software to another ($p < 0.001$) with the absolute CSFV showing the highest variations followed by GMV, however, the corresponding normative ratios did not show such a variation. To the best of our knowledge, no single study performed such a comparison among these three software, with all prior work focusing on subcortical gray matter segmentation or hippocampal segmentation, the few work addressing the global volumetric measures used either phantoms or AD cases for assessment and used SPM itself instead of CAT12 box and finally held inter-software comparison using absolute global volumetric measures instead of normative ratios as in Fellhauer et al. [9], or Palumbo et al. [10].

Although our study targeted population ≥ 15 years old, we found that both age and gender modulated the absolute global volumetric measures (p -value < 0.001). The effect of age was more pronounced in males possibly because the absolute male volumes were constantly higher than the female one; hence, it was easier to capture aging changes in it. Nonetheless, our current data model could not be used to extrapolate such a relation in full due to lack of Egyptian's below 15 years of age and/or longitudinal data model as in Ge et al. [11].

The effect of age on absolute volumetric measures echoed through the corresponding normative ratios generated by all software. Yet, FSL-FAST-generated absolute volumes and normative ratios best demonstrated the effect of age and gender on global volumetric measures except for: (1) absolute and normative ratio of GMV, in which FreeSurfer constantly outperformed both FSL-FAST and SPM-CAT12 yielding the largest mean GMV with the least degree of variance. Such a finding can be attributed to the superiority of FreeSurfer algorithm in segmenting gray matter as well as the aforementioned conundrums of segmenting CSF and WM masks. (2) Absolute CSFV in males; possibly because the larger cortical gray matter volume in males could imply a smaller sulcal CSFV, hence, the gap between the true and mathematically summated CSFV is reduced.

The effect of gender however was only captured in the normative ratio generated by the software with the largest absolute volume and least degree of variance of

a certain measure, which are FreeSurfer for GMV and FAST for WMV and CSFV.

The mean and standard deviation of absolute and relative GMV, WMV, CSFV and TIV in Egyptians were calculated for each gender and software; however, edge to edge comparison of the Egyptian global volumetric data to other ethno-racial data as in Guttman et al. [12], on Americans could not be done because of the differences in data categorization, e.g., differentiating data by gender, utilized age increments and occasionally the longitudinal data collection pattern, for instance, Guttman et al. [12], did not differentiate between male and female volumes and moreover, they summated the volumetric data of all subjects < 40 years in one category then used a 10 years incremental step in population > 40).

According to the results of our study, FSL-FAST was considered to be overall the most efficient program to generate global brain volumetric results—unless the GMV is the primary study objective—because it consumed the least processing time and its normative ratios did not vary significantly from that of SPM-CAT12 or FreeSurfer, on the contrary its absolute and relative volumes effectively captured gender and age effect on global brain volume except for the gray matter. Moreover, if subcortical gray matter volume is required, combining FSL-FAST and FSL-FIRST will consume lesser time than the time needed by SPM-CAT12 to generate global volumetric measures alone with comparable subcortical segmentation results to FreeSurfer as in Mulder et al. [13]. It is true it did require some in-house code development to effectively run multiple-subject analysis to extract its numeric output in a user-friendly format, yet the extracted data were to the point and easily interpretable by non-specialized user.

Although global and subcortical brain volumetric measures (including total hippocampal volume) produced by FSL-FAST and FSL-FIRST could suffice in many clinical settings, for instance follow-up of global brain atrophy in MS or diagnosis of mesial temporal sclerosis, it will not suffice in a vast set of clinical applications like differential diagnosis of neurodegenerative diseases and usually another set of structural (like entorhinal cortical thickness in differentiating AD from FTD) or pathological [(such as white matter lesion load (WMLL) or T1 black hole volume in MS)] volumetric measures will be required [14].

It is true that some of these structural measures can still be obtained using non-robust FSL ROI-based analysis tools necessitating lots of manual editing, yet the more robust atlas/ROI-based analysis conducted by SPM-CAT12 in which one can select the required subset of structural measures or FreeSurfer in which a forced subsegmental analysis using several atlases as

well as white matter pathological volumes will be generated with minor usually insignificant error rate [14, 15].

In fact, white matter pathological volumetric results could be derived from another FSL tool (FSL-BIANCA) or SPM tool (SPM-LST), yet both are totally separate analysis processes requiring lots of manual editing and revision and will only evaluate for white matter hyper- and not hypo-intensities.

Conclusions

The efficiency of certain brain volumetric analysis software is better defined by the clinical objective, whenever a rapid check of global and subcortical brain volumetric measures is needed, FSL-FAST and FSL-FIRST would be the most cost-effective option; on the contrary, if a comprehensive analysis is needed, FreeSurfer is rather a fault-proof choice supported by results; however, further evaluation of SPM-CAT12 subsegmental brain volumetric measures is still needed especially that it has the most friendly GUI and output file structure as well as customizable volumetric analysis.

Abbreviations

| | |
|---------|---|
| ANOVA | Analysis of variance |
| BET | Brain extraction tool |
| CAT12 | Computational anatomy toolbox |
| CSF | Cerebrospinal fluid |
| CSFV | Cerebrospinal fluid volume |
| DICOM | Digital imaging and communications in medicine |
| DWI | Diffusion-weighted images |
| eTIV | Estimated TIV |
| FAST | Fmrib's automated segmentation tool |
| FIRST | FMRIB's integrated registration and segmentation tool |
| FLAIR | Fluid-attenuated inversion recovery |
| FSL | FMRIB's software library |
| GM | Gray matter |
| GMV | Gray matter volume |
| MNI | Montreal neurological institute |
| MP-RAGE | Magnetization-prepared rapid acquisition with gradient-echo |
| MRI | Magnetic resonance imaging |
| NIFTI | Neuroimaging informatics technology initiative |
| SD | Standard deviation |
| SPM | Statistical parametric mapping |
| WM | White matter |
| WMV | White matter volume |

Acknowledgements

Not applicable.

Author contributions

All authors read and approved the final manuscript. Experimental design: AAAH, AOA, AAEB. Data interpretation: AAAH, HMA, LMAMM, MA. Statistical analysis: HMA. Neurological examination: HMA, LMAMM. Writing and revision: AAAH, AOA, HMA, AAEB. The manuscript has been read and approved by all the authors, and the requirements for authorship as stated earlier in this document have been met, and each author believes that the manuscript represents honest work.

Funding

Not applicable (no funding received for this study).

Availability of data and materials

All the datasets used and analyzed during this study are available with the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

All procedures performed in the studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This study was approved by the research ethics committee of the Radiology department of the Faculty of medicine Cairo University on 4/2019. Informed consent was obtained from all individual participants included in the study.

Consent for publication

All participants included in this study gave an informed consent to publish the data contained in this study.

Competing interests

No potential conflict of interest was reported by the authors.

Received: 15 March 2023 Accepted: 19 August 2023

Published online: 28 August 2023

References

1. Abrigo J, Shi L, Luo Y, Chen Q, Chu WCW, Mok VCT, Alzheimer's Disease Neuroimaging Initiative (2019) Standardization of hippocampus volumetry using automated brain structure volumetry tool for an initial Alzheimer's disease imaging biomarker. *Acta Radiol* (Stockholm, Sweden: 1987) 60(6):769–776. <https://doi.org/10.1177/0284185118795327>
2. Raji A, Ostwaldt AC, Opfer R, Suppa P, Spies L, Winkler G (2018) MRI-Based brain volumetry at a single time point complements clinical evaluation of patients with multiple sclerosis in an outpatient setting. *Front Neurol* 9:545. <https://doi.org/10.3389/fneur.2018.00545>
3. Akudjedu TN, Nabulsi L, Makelyte M, Scanlon C, Hehir S, Casey H, Ambati S, Kenney J, O'Donoghue S, McDermott E, Kilmartin L, Dockery P, McDonald C, Hallahan B, Cannon DM (2018) A comparative study of segmentation techniques for the quantification of brain subcortical volume. *Brain Imaging Behav* 12(6):1678–1695. <https://doi.org/10.1007/s11682-018-9835-y>
4. Katuwal GJ, Baum SA, Cahill ND, Dougherty CC, Evans E, Evans DW, Moore GJ, Michael AM (2016) Inter-method discrepancies in brain volume estimation may drive inconsistent findings in Autism. *Front Neurosci* 10:439. <https://doi.org/10.3389/fnins.2016.00439>
5. Guo C, Ferreira D, Fink K, Westman E, Granberg T (2019) Repeatability and reproducibility of FreeSurfer, FSL-SIENAX and SPM brain volumetric measurements and the effect of lesion filling in multiple sclerosis. *Eur Radiol* 29(3):1355–1364. <https://doi.org/10.1007/s00330-018-5710-x>
6. Backhausen LL, Herting MM, Buse J, Roessner V, Smolka MN, Vetter NC (2016) Quality control of structural MRI images applied using FreeSurfer—a hands-on workflow to rate motion artifacts. *Front Neurosci* 10:558. <https://doi.org/10.3389/fnins.2016.00558>
7. McCarthy CS, Ramprasad A, Thompson C, Botti JA, Coman IL, Kates WR (2015) A comparison of FreeSurfer-generated data with and without manual intervention. *Front Neurosci* 9:379. <https://doi.org/10.3389/fnins.2015.00379>
8. Beelen C, Phan TV, Wouters J, Ghesquière P, Vandermosten M (2020) Investigating the added value of FreeSurfer's manual editing procedure for the study of the reading network in a pediatric population. *Front Hum Neurosci* 14:143. <https://doi.org/10.3389/fnhum.2020.00143>
9. Fellhauer I, Zöllner FG, Schröder J, Degen C, Kong L, Essig M, Thomann PA, Schad LR (2015) Comparison of automated brain segmentation using a brain phantom and patients with early Alzheimer's dementia or mild cognitive impairment. *Psychiatry Res* 233(3):299–305. <https://doi.org/10.1016/j.pscychres.2015.07.011>

10. Palumbo L, Bosco P, Fantacci ME, Ferrari E, Oliva P, Spera G, Retico A (2019) Evaluation of the intra- and inter-method agreement of brain MRI segmentation software packages: a comparison between SPM12 and FreeSurfer v6.0. *Phys Med PM Int J Devot Appl Phys Med Biol Off J Ital Assoc Biomed Phys (AIFB)* 64:261–272. <https://doi.org/10.1016/j.ejmp.2019.07.016>
11. Courchesne E, Chisum HJ, Townsend J, Cowles A, Covington J, Egaas B, Harwood M, Hinds S, Press GA (2000) Normal brain development and aging: quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology* 216(3):672–682. <https://doi.org/10.1148/radiology.216.3.r00au37672>
12. Guttman CR, Jolesz FA, Kikinis R, Killiany RJ, Moss MB, Sandor T, Albert MS (1998) White matter changes with normal aging. *Neurology* 50(4):972–978. <https://doi.org/10.1212/wnl.50.4.972>
13. Mulder ER, de Jong RA, Knol DL, van Schijndel RA, Cover KS, Visser PJ, Barkhof F, Vrenken H, Initiative ADN (2014) Hippocampal volume change measurement: quantitative assessment of the reproducibility of expert manual outlining and the automated methods FreeSurfer and FIRST. *Neuroimage* 92:169–181. <https://doi.org/10.1016/j.neuroimage.2014.01.058>
14. Bartos A, Gregus D, Ibrahim I, Tintëra J (2019) Brain volumes and their ratios in Alzheimer’s disease on magnetic resonance imaging segmented using FreeSurfer 6.0. *Psychiatry Res Neuroimaging* 287:70–74. <https://doi.org/10.1016/j.psychres.2019.01.014>
15. Grimm O, Pohlack S, Cacciaglia R, Winkelmann T, Plichta MM, Demirakca T, Flor H (2015) Amygdalar and hippocampal volume: a comparison between manual segmentation, FreeSurfer and VBM. *J Neurosci Methods* 253:254–261. <https://doi.org/10.1016/j.jneumeth.2015.05.024>

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
